

Few words strike fear in the hearts of school age children more than the word “test”. I have seen the same look of worry on the faces of my fifth grade students every time I mention that my Spanish classes will include two tests during the school year. While I use the word “test” for dramatic effect on that first day in order to get everyone’s attention, students soon learn that these two assessments are completely different from anything that they have seen before. They have no need to panic.

A Test That Isn’t Torture: A Field-Tested Performance-Based Assessment

Mark Eastburn

Changing the paradigm

I have experimented with traditional tests, similar to the ones that I took as a high school Spanish student, but my apprehensive early language learners gave a mediocre performance that didn’t seem to reflect their actual skills. Reviewing their test results often left children feeling that, despite their hard work, they weren’t learning Spanish. Their confidence was shaken. Also, these tests did not measure the interpersonal and presentational modes of communication at all, and the measure of interpretive communicative ability they provided was limited to a relatively low number of vocabulary words and phrases. Still needing evidence of student progress for parents and administrators, I realized the need to break the old testing paradigm and discover a new way of demonstrating student language acquisition.

The performance-based evaluation that I now give to students every October and June is a wonderful opportunity to measure and showcase each child’s ability to read, listen, speak and exchange information in Spanish. This assessment also gives me a tremendous amount of evidence to document each student’s progress over the course of the year. As you will see, the performance assessment makes it possible to measure students at all ability levels at the same time. This is particularly helpful in Princeton, New Jersey, since I often have heritage speakers sitting next to first year students. The performance assessment creates an experience that motivates children, rapidly overcomes their test anxiety, and (could it be?) seems fun!



The Johnson Park School Fifth Grade Assessment



Johnson Park School is a K-5 public elementary school located in Princeton, NJ. We have a diverse student body, with approximately 15% of the children coming from Spanish-speaking households and another substantial minority that is composed of heritage speakers of more than 20 languages. We receive new students at every grade level each year, so we frequently face the challenge of creating Spanish classes that are simultaneously meaningful to heritage speakers and students of very limited proficiency. We teach our Spanish classes from kindergarten through fifth grade using content and culture rich thematic units to help us differentiate the learning effectively. Our assessment, therefore, needs to touch on all of the thematic units that we study in fifth grade in order to ensure an accurate measure of the important language functions taught throughout the year.

The unifying theme for these performance-based activities is exploration of a museum. The fifth grade Spanish curriculum in our district consists of five thematic units; the culminating assessment task for each unit represents a dif-

ferent part of our museum. My main goal is to measure each student's abilities in all three modes of communication as defined by the ACTFL Performance Guidelines for K-12 Learners: the interpretive mode, the presentational mode, and the interpersonal mode (Abbott, Boyles, & Miles, 1998). As a result, the test is divided into three parts.



Interpretive Assessment

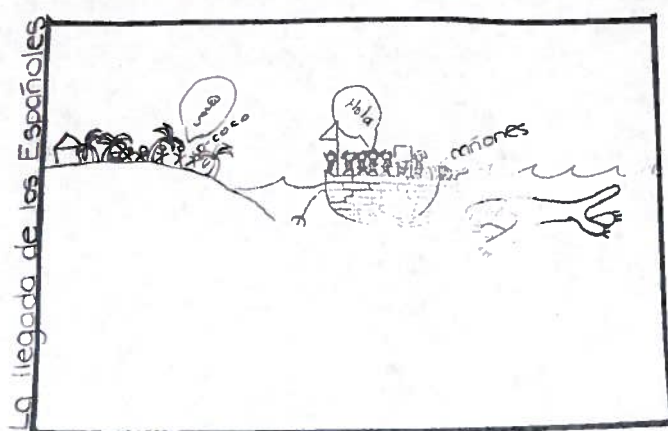
Part one measures the ability of students to interpret language that they read and hear (interpretive mode). In this case, I ask students to draw a blueprint of our fictional museum based on a description that I read aloud and allow them to reference for themselves. I have written three different paragraphs that describe three different rooms in the building. Students receive a blank sheet of paper that has three rectangles. Each blank rectangle represents a different exhibit in the museum. After I give the initial instructions to place the task in context, I hand a copy of the complete description to each student and read aloud the first paragraph. Students then have fifteen minutes to draw everything that they can identify in this first exhibit. Of course, since this is not art class, I emphasize that pictures do not need to be beautiful, only simple and clear. Very basic depictions of objects and actions are perfectly acceptable, as can be seen in figure 1. I also allow students to write notes on the test sheet in order to clarify drawings that they think may not be clear to me. I repeat this process for the two remaining rooms and collect all of the papers when students are finished. Since I usually teach for thirty-minute periods, the interpretive assessment takes approximately two class periods to complete.

I designed the detailed descriptions to differentiate among different levels of language comprehension. For example, one room in our museum focuses on our thematic unit: The Discovery and Conquest of the New World by Spain in the 1500s. The paragraph-long description of the exhibit contains simple, easily identifiable words like *isla* (island) and *barco* (boat) that a Novice-Low student would recognize, but it also includes strings of sentences that would only be comprehensible to a Novice-High learner: "*Hay siete marineros en la cubierta del barco. Cuatro de éstos hombres tienen sus manos levantadas. Ellos están saludando a los indígenas en español!*" (There are seven sailors on the deck of the boat. Four of these men have their hands raised. They are greeting

the indigenous people in Spanish.) Since I am not grading artistic prowess, I would accept seven stick figures on a boat, four with their hands up, and a word bubble that contained the word *hola* (hello) as representing all of the relevant information contained in this passage. Student work is then graded on a five point rubric, where a "1" represents very little or no comprehension of the material presented, a "5" represents a nearly complete accounting of all details, and a "4" represents comprehension that would be expected of a typical listener/reader at the Novice-High level of the American Council on the Teaching of Foreign Language scale. I marked the example shown as a "5" since practically every detail from the description is included, even the passage that states that the mast in the front with the triangular flag is higher than the mast in the back with the square flag. A "4" would have the right number of sailors in the boat and the correct number of indigenous people and other objects on the island, but might lack the very specific details like the difference in mast height that this example shows. The comprehension that would be expected of Novice-Mid students and Novice-Low students correspond respectively to the "3" and "2" on this rubric. A student who has drawn isolated words such as an island, a crocodile, and a boat would receive a "2," while a boat with sailors in it and trees on the island (which are words that appear in contiguous sentences) would receive a "3."

Figure 1

Interpretive Drawing



Presentational Assessment

The presentational part of this performance-based assessment measures each student's ability to present information to an audience with "no direct opportunity for the active negotiation of meaning" (Abbott, Boyles, & Miles, 1998, p. 4). Fifth grade classes learn about art from the Spanish-speaking world. In part two of the assessment, I tell students that our museum is expanding to include an art exhibit. The museum



choose their technique after one of the artists that we have studied, such as Picasso or Botero. The test takers have complete freedom to include any objects, people, actions, landscapes, and weather phenomena in their drawings that are appropriate for school and they can describe in Spanish. After twenty-five minutes I collect their papers to discourage the addition of material or outside help. As in part one, I emphasize that I will not be considering their artistic ability in my grading. I do not expect them to recreate a masterpiece only that they can describe their work in Spanish. In fact, this past year I had two students who did not draw anything and then proceeded to give wonderful descriptions of everything that wasn't on their paper!

On the second day, I select students at random to tell me about their creation. Students do not rehearse their presentations, but they may refer to their drawings and any notes from the prior class. Based on student presentations of their drawings over the last three years, I have developed a holistic rubric that defines the performances of every single student neatly and unambiguously in one of the five categories. This rubric was adapted from the descriptions of different language ability levels in the New Jersey World Language Core Curriculum Content Standards and the College Entrance Examination Board.

5 Exceeds Expectations Novice-High

Spanish can be understood easily. The vocabulary is varied and provides a great deal of information. The speaker demonstrates confidence when communicating in Spanish, using strings of sentences that have continuity. Speaker does not need any assistance from the teacher when presenting.

4 Meets Expectations Novice-High

Speaker presents information in strings of sentences with frequent repetition of vocabulary and/or structures (e.g. "*La niña tiene una camisa. La niña tiene dos brazos. La niña tiene una falda.*"). Student speaks with reasonable ease without needing assistance or encouragement from the teacher.

3 Meets Expectations Novice-Mid

Demonstrates basic ability to communicate in Spanish in a guided format and familiar context. Speaker can put words into simple phrases or sentences using familiar vocabulary and memorized structures. Spanish can be understood with

some effort. Speaker may need some assistance with vocabulary and direction of the presentation from the teacher.

2 Meets Expectations Novice-Low

Demonstrates a limited ability to communicate in Spanish. Speaker produces lists of isolated words related to the drawing, but generally depends on the teacher's support to communicate any additional information in phrases or sentences. Spanish may be halting and difficult to understand.

1 Does Not Meet Expectations Beginner


Description is in English or only consists of very few words in Spanish. Even with teacher support, speaker can produce very little Spanish in this context.

Interpersonal Assessment

The interpersonal assessment involves impromptu conversations and active negotiation of meaning among the participants. While interpersonal communication between students and the teacher, students and heritage speakers, and students and community members is a central focus of world language classes in Princeton, I have still had to revise this portion of the assessment several times due to the sheer challenge of getting Novice level fifth graders to talk to each other and maintain conversation in Spanish! Among the numerous activities that I have attempted in order to encourage students to talk to each other, the most effective has been an information gap activity in which students assume the roles of servers in the museum's restaurant. This setting corresponds to our fifth grade thematic unit on food and places where people eat. In this scenario, four students acting as servers come together to discuss what is happening at the tables of patrons that they can see from a central vantage point. While pairs of students speaking to each other or individual students engaged in a conversation with the teacher would work well just as well for Part Three, I use groups of four in order to measure a greater number of students simultaneously during my thirty minute class periods. I currently represent the different tables with nine simple drawings. In the future, I plan to have actual photographs for each scenario.



Some of the actions and objects that appear in these scenes are repeated in multiple images, and others are unique to



each table. Students have been working with the vocabulary, gestures, and phrases over the course of learning about food and restaurants. For example, three pictures have a man asking for the check with his index finger and thumb clasped as if he were holding a pen. I place all nine drawings labeled with a letter on the board and ask four students to the front of the class. Each student secretly gets a copy of one of the pictures that appears on the board.

I then explain to students that they are servers who are working in the museum restaurant. Each server needs to figure out what is going on in the other servers' areas (the pictures that they cannot see). They will need to accomplish this task by asking and answering questions that they create, and they may reference the nine pictures that they can see on the board. They may not, however, ask which letter label another child has. One student may ask another, "*¿Hay un hombre que necesita la cuenta?*" (Is there a man who needs the check?), to which the other student may reply, "*¡Sí!*" (Yes) or maybe, "*No, pero un hombre está comiendo*" (No, but a man is eating). Each of the four students is given multiple opportunities to ask and answer questions, and they are graded both on the complexity of questions that they ask and the answers that they give. In this case detailed questions, follow up questions, and intricate answers will be awarded a "5" for far exceeding expectations for fifth grade Spanish; well-phrased questions and answers that show some detail will receive a "4" for meeting expectations of a Novice-High student; and simple questions and answers will earn a "3" for a Novice-Mid student. Questions and answers that are in English or incomprehensible are given a "1". The "2" on my five-point rubric documented the ability of certain low-proficiency students to benefit from exposure to the language of higher-proficiency students in their group and mimic others questions and answers.

Results

The three paragraphs that form the interpretive part of the assessment were over 500 words. I wrote these paragraphs so that very few words are repeated in the description. In its current form, the test measures comprehension in a meaningful context and includes nouns, verbs in different conjugations, and description. Students showed no signs of burnout or exhaustion after 45 minutes. The assessment can easily be expanded as I decide to include additional vocabulary and language structures. I was able to rapidly grade the student drawings of the museum exhibits. I looked at each picture for only a minute to determine what the student was able to comprehend. I completed reviewing and grading all fifty-three students' work in less than an hour and a half.

In the presentational assessment, many students created speeches that lasted three minutes or more. I allowed them to speak until they felt satisfied that they had done their best. If students were hesitant, I prompted them with questions in order to get them talking, but few needed much help. The open-ended nature of the assessment allowed students to show what they could do without limits on their production. Many of these children scored above the "3" that I was expecting for Novice-Mid students, and a rather large number exceeded my expectations for Novice-High speakers. I suspect that this achievement was primarily due to students' freedom to choose their own topics and ability to work from a picture that they created. This familiarity lowered their anxiety and enabled them to speak from a much more comfortable mental state.

The interpersonal assessment, while still the most difficult portion of the test to administer and assess, was also successful in encouraging student communication. Students of lower proficiency were put at ease when other children in the group modeled appropriate questions and answers. No one froze when it was their time to speak, in part because I let my most confident students begin the questioning session each time.

Parts Two and Three were graded in real-time; that is to say that I was able to decide a number on the rubric for each student at the conclusion of his or her participation in the activity. This was another advantage to this method of assessment, since I did not need to set aside additional time to grade student work outside of class. I did, however, videotape these activities in order to review each child's performance and determine if my previous decision was the most accurate. After a careful analysis, I realized that I did not need to change any of the numbers that I had originally assigned to students.

I give the same assessment to students at the beginning and the end of the year to demonstrate their progress in Spanish. Students are not allowed to see any part of the test in the intervening time, and yet most make substantial progress on the rubrics from October to June. The average progress made by each student was 1.7 points on the five-point scale, so a typical student who scored a 2.3 on each part of the assessment in October would finish the year with a 4.0 on each rubric in June.

"I thought that this test was good because it was easier and had less pressure." - Johnson Park student

Student Feedback

After completing the test this year, I asked all of my students to reflect on the nature of this assessment in comparison to the tests that they take in other subjects. Student opinion was overwhelmingly in favor of this type of examination. Joshan Bajaj, a student at Johnson Park School for the past two years, summed up the opinion of the majority: "The test was a lot of fun. I hope we will take this test again soon. The only thing is that the test could be a little harder." Of course, this test is actually more comprehensive and rigorous than many traditional tests. It does seem easier to most students, primarily because it serves as an opportunity for students to demonstrate what they do know and can do, instead of highlighting everything that they cannot do in Spanish.

Other students responded anonymously with reactions of relief and brutal honesty, "I don't like tests, but this was better than most tests", and with interesting Spanish phrases "*Me mucho me gusta dibujitos. No me gusta escribe-ing*" (I like the drawings. I don't like writing). This test was a tremendously positive experience, in stark contrast to the traditional tests that students dread in other subjects. A few students even asked for more after we were finished!

Tailoring the Assessment

The Johnson Park Fifth Grade Assessment may be adapted to your classes. The unifying theme of a museum can work with almost any unit or chapter since there are museums for most everything. When you complete a content-rich thematic unit or even a chapter in a textbook that includes a great deal of vocabulary, asking children to draw a picture from a description effectively and quickly assesses your students' ability to interpret language. Your description can be as simple or as complex as the language ability of your students, and differentiation of assessment can be established by including both simple and complex passages in one description. Student presentations with only a short opportunity for preparation, as we saw in the presentation assessment, may similarly be adapted to any theme, and the images in our information-gap activity for the interpersonal assessment could be recreated to include any additional words, settings, or concepts that you choose. While such activities do take time to prepare, the information that you gather on the progress of your students will serve as powerful evidence of their success in your class. This data speaks loudly to parents and administrators when we are advocating for our world language programs.



Conclusion

Of course, I could never have created this performance-based assessment entirely on my own. I owe a huge debt of gratitude to our principal, Dr. Robert Ginsberg, for having the vision to place world language assessment in our school's goals during the 2003-2004 and 2004-2005 academic years. In addition, I have benefited tremendously from the guidance and support of our former world language supervisor, Barbara Greenfeldt, our current world language supervisor, Priscilla Russel, and Greg Duncan of InterPrep, Inc. Thank you to all.

References

- College Entrance Examination Board. (1996). *Articulation and Achievement Connection Standards, Performance and Assessment in Foreign Language*. New York: College Board
- Abbott, M., Boyles, P., & Miles, J. (Eds.). (1998). *ACTFL Performance Guidelines for K-12 Learners*. Yonkers, NY: ACTFL.



Mark Eastburn is a Spanish teacher for grades 3, 4, and 5 at Johnson Park School in Princeton, NJ. Prior to five years as a Spanish teacher, Mark worked as an Agroforestry Volunteer for the U.S. Peace Corps in Las Huacas de Quije, a small village in the mountains of central Panama. For three weeks in 2003, he traveled to Bolivia and Argentina as a "Teacher in the Field" for the South American Low Level Jet Experiment (SALLJEX). In addition to his teaching duties, Señor Eastburn is writing on his first thematic unit book on endangered species in Latin America and culturally authentic crafts from Central and South America.